

An SGML/XML Roadmap

Presentation to ISO TC184/SC4
by Daniel Rivers-Moore, RivCom
Orlando, FL, 4 February 1998

What is SGML?

- Standard Generalized Markup Language
- ISO 8879 (1986 + TC 1997)
- Abstract syntax for markup languages
- Markup = metadata
- Typically – but not necessarily – used for documents (information structured for human consumption)

Tagging for document structure

```
<!ELEMENT book (frontmatter, mainmatter,  
endmatter)>  
  
<!ELEMENT frontmatter (title, author,  
copyright, preface?, dedication?, toc?)>  
  
<!ELEMENT mainmatter (chapter+)>  
  
<!ELEMENT preface (title, para+)>  
  
<!ELEMENT chapter (title, para+)>  
  
<!ELEMENT title (#PCDATA)>  
  
<!ELEMENT para (#PCDATA)>  
  
...
```

Tagging for document structure

```
<book><frontmatter><title>Captain Corelli's  
Mandolin</title><author>Louis de Bernières  
</author><copyright>1994 by Louis de  
Bernières</copyright><dedication>To my  
mother and father</dedication></frontmatter>  
<mainmatter><chapter><title>Dr Iannis  
Commences His History and is Frustrated  
</title><para>Dr Iannis had enjoyed a  
satisfactory day ...  
...  
</book>
```

What is HTML?

- An SGML application
- Tagging for presentation and behaviour in a Web browser
- Validity of the SGML not enforced by the browsers

Tagging for presentation/behaviour

```
<!ELEMENT HTML (HEAD?, BODY)>
<!ELEMENT BODY (blockstuff | inlinestuff |
formstuff)*>
<!ELEMENT blockstuff (heading | P | LIST)>
<!ELEMENT inlinestuff (B | I | A)>
<!ELEMENT formstuff (checkbox | button |
textbox)>
<!ELEMENT HEADING ( H1 | H2 | H3 | H4 | H5 |
H6 | H7 )>
<!ELEMENT LIST (LI+)>
```

Tagging for presentation/behaviour

```
<!ATTLIST A  
    NAME CDATA #IMPLIED  
    HREF CDATA #IMPLIED  
    -- The HREF attribute is the URL of the  
    other end of the hyperlink --  
>
```

What is XML?

- eXtensible Markup Language
- Simplified SGML for use on the Web
- Not tied to a specific tag set
- Not tied to specific semantics
- Allows structured data to be transmitted over the Web

Tagging for structured data

```
<meeting name="Liaison Plenary">
  <date>4 February, 1998</date>
  <start_time>8.00 am</start_time>
  <end_time>10.00 am</end_time>
</meeting>
<liaison from="JTC1/WG4" status="proposed">
  <person personID="ek"/>
</liaison>
<liaison from="TC184/SC5/WG4">
  <person personID="hm"/>
</liaison>
```

Tagging for structured data

```
<person ID="ek">  
  <name>Eliot Kimber</name>  
  <company>Isogen International</company>  
  <country>USA</country>  
</person>  
<person ID="hm">  
  <name>Helium Mak</name>  
  <company>CNRC</company>  
  <country>Canada</country>  
</person>
```

Some Applications of XML

- CDF (Channel Definition Format)
- OSD (Open Software Definition)
- PNG (Portable Network Graphic)
- RDF (Resource Definition Format)
- XML-EDI
- ... and more

An SGML/XML Roadmap

Markup

SGML

XML

Links

HyTime

XLL

Presentation/Behaviour

DSSSL

XSL

Schema Definition

?

XML-Data?

